**Title: Validity of the Waterlow screening tool and risks for pressure injury in acute care.**

Joan Webster, Royal Brisbane and Women's Hospital, University of Queensland, Queensland University of Technology, and Griffith University.

Nicole Gavin, Royal Brisbane and Women's Hospital

Catherine Nicholas, Royal Brisbane and Women's Hospital

Kerrie Coleman, Royal Brisbane and Women's Hospital

Glenn Gardner, Royal Brisbane and Women's Hospital and Queensland University of Technology

**Correspondence author.**

Prof Joan Webster

Nursing Director, Research

Level 2, Centre for Clinical Nursing

Royal Brisbane and Women's Hospital

Herston QLD 4029

AUSTRALIA

T: +61 7 3636 8590 F: +61 7 3636 5832 E-mail joan_webster@health.qld.gov.au

**ABSTRACT**

**Objectives:** To assess the validity of the Waterlow screening tool in a cohort of internal medicine patients and to identify factors contributing to pressure injury.

**Design:** Longitudinal cohort study

**Setting:** A tertiary hospital in Brisbane, Australia

**Participants:** 274 patients admitted through the Emergency Department or outpatient clinics and expected to remain in hospital for at least three days were included in the study. The mean age was 65.3 years.

**Interventions:** Patients were screened on admission using the Waterlow screening tool. Every second day, their pressure ulcer status was monitored and recorded.

**Main outcome measures:** Pressure ulcer incidence

**Results:** Fifteen participants (5.5%) had an existing pressure ulcer and a further 12 (4.4%) developed a pressure ulcer during their hospital stay. Sensitivity of the Waterlow scale was 0.67, (95% CI: 0.35 to 0.88); specificity 0.79, (95% CI: 0.73 to 0.85); PPV 0.13, (95% CI: 0.07 to 0.24); NPV 0.98, (95% CI: 0.94 to 0.99).

**Conclusion:** This study provides further evidence of the poor predictive validity of the Waterlow scale. A suitably powered randomised controlled trial is urgently needed to provide definitive evidence about the usefulness of the Waterlow scale compared with other screening tools and with clinical judgement.

**KEY WORDS** Cohort studies; Predictive value of tests; Pressure ulcer.

**KEY PHRASES**

- Pressure ulcer incidence is a key indicator of nursing care.
- The Waterlow scale is widely used to predict which patients may develop a pressure ulcer but its validity and reliability varies between settings.
- In acute settings, clinical judgement should be combined with Waterlow scale assessment to avoid over classification of patients into a high risk category, which has implications for scarce resources.
- Possible reasons for the limited the validity of the Waterlow scale in acute hospital are provided.

## INTRODUCTION

Pressure ulcers continue to represent a significant clinical problem in health care facilities and are associated with considerable morbidity and mortality (Kuwahara et al, 2005; Landi et al, 2007). Costs to the healthcare system in Australia have been estimated at (AUD) $285 M per annum, representing a considerable financial burden on hospital budgets (Graves et al, 2005). In acute settings pressure ulcer prevalence rates vary between 2.9 and 32.4% (Lapsley and Vogels, 1996; Keong et al, 2004; Nixon et al, 2007) depending on the type of patients included in the study cohort and the criteria used to define a pressure ulcer. In our own hospital, the average pressure ulcer prevalence rate is 12.7% but this rate varies between specialty areas from 3.4% to 21.4%, with lowest rates in women's and newborn services and highest in critical care services.

One of the long-standing approaches to pressure ulcer prevention is screening for factors thought to be associated with pressure injury. Many tools are available for this purpose but the Waterlow screening tool is one of the most widely used. It was developed over 20 years ago from an initial survey in an elderly care ward and a subsequent survey in acute areas of the hospital (Waterlow, 1985). Although early testing appeared positive, more recent studies have identified weaknesses in the tool, such as its poor predictive validity; particularly a tendency for the tool to overestimate the number of patients at risk (Pancorbo-Hidalgo et al, 2006). One suggestion is that locally-determined risk factors need to be incorporated into risk assessment tools to improve their effectiveness (Papanikolaou et al, 2007).

At the study hospital the Waterlow screening tool has been used for a number of years; despite this, no lasting reduction in the prevalence of pressure ulcers has been observed. This may be because the instrument is not effective in an acute care setting, where risk factors may differ to those in community facilities. We currently have no information about the risks associated with pressure injury in this institution, or if these risks vary between specific units of the hospital. It is important that this preliminary work is undertaken to guide any quality improvement pressure-ulcer prevention interventions that may be implemented in the future.

**OBJECTIVES:**

The objectives of the study were to assess the validity of the Waterlow instrument in a cohort in internal medicine patients and to identify factors contributing to pressure injury.

**METHODS**

**Research design**

We used prospective cohort design. This is an observational study that follows a group of people (a cohort) over a period of time to investigate the effect of an intervention; in this case, administering the Waterlow scale.

**Population**

Patients admitted to any internal medicine ward at the Royal Brisbane and Women's Hospital (RBWH) were eligible for inclusion. Screening for pressure ulcer risk, using the Waterlow screening tool is routine procedure at the hospital. Moreover, as audits of pressure ulcer prevalence are regularly conduct, and as no additional burden was placed on participants, the Institutional Ethics Committee deemed that no formal patient consent was required.

**Instuments**

1. The Waterlow pressure ulcer screening tool (Waterlow, 1985) was chosen for the study because it is the one recommended for use at our hospital. The scale consists of eight items: build/weight for height, visual assessment of the skin type, age and sex, continence, mobility, a measure of malnurition, and a 'special risk' factor (including tissue malnutrition, neurological deficits and major surgery or trauma). Highest and lowest scores vary for each item and a score of 16 or above is the generally accepted cut-off point for at-risk patients.

2. A survey tool, including demographic information and a list of risk factors associated with pressure ulcer development compiled from the literature, was used for data collection. The list contained items relating to diagnosis on admission, any existing comorbidities, physical and mental status, biochemistry, and medication use.

**Procedure**

Patients expected to remain in hospital for at least three days were screened, using the Waterlow screening tool, by a research nurse trained in its use. Baseline demographics including age, gender, principal diagnosis, co-morbidities, and medication use were recorded on admission. Presence of existing pressure ulcers were also documented. Every second day patients were reviewed until the development of a pressure ulcer or until discharge. Presence of any risk factor was recorded using a Yes/No response to the list of risk factors.

Research nurses assigned a 'stage' to all pressure ulcers herself/himself using the National Pressure Ulcer Advisory Panel's updated pressure ulcer staging system

(Black et al, 2007). The seven research nurses, who participated in data collection, undertook inter-rater reliability testing of 1) staging pressure ulcers, using four multiple choice questions and photographs of pressure ulcers, and 2) scoring the Waterlow screening tool, using a series of case studies. For the Waterlow analysis, scores were categorised as: 0 – 10 low risk, >10 at risk, >15 high risk, and 20+ very high risk. Staff in participating wards were advised of the study but no changes were made to routine care.

**Analysis**

Data were entered and analysed using SPSS version 16.0. Baseline demographic and risk factor characteristics are reported as frequencies or means and standard deviations (SDs). The statistical analysis focused on three specific outcomes: pressure ulcer incidence, validity of the Waterlow screening tool and the association between risk factors and pressure ulcer development.

*Pressure ulcer incidence*

We calculated the pressure ulcer incidence rate as the number of new pressure ulcers occuring as a proportion of the total population of patients included in the study.

*Validity of the Waterlow*

Sensitivity (the proportion of patients correctly classified as high risk of developing a pressure ulcer), specificity (the proportion of patients correctly classified as low risk of developing a pressure ulcer), positive predictive values (PPV the proportion of those classified as high risk who developed a pressure ulcer) and negative predictive values (NPV proportion of those classified as low risk who did not develop a pressure ulcer) were calculated using a two by two table. For the Waterlow tool analysis, the recommended cut-off point of > 15 was used.

*Risk factor identification*

We calculated the crude odds ratios and their 95% confidence intervals (CI) for the proportion of patients who developed a pressure ulcer for each of the demographic variables and each of the risk factors we measured. The patient was the unit of analysis, irrespective of the number of pressure ulcers.

*Inter-rater reliability*

Inter-rater agreement was assessed using the Intraclass Correlation Coefficient (ICC's) statistic with confidence intervals of 95%. The ICC measures how much of the total variance of scores can be attributed to differences between participants (Bravo and Potvin, 1991). Poor correlation and systemic score differences result in reduced values. ICC values range from 0 to 1; values of 0.7 and over are considered to indicate 'substantial agreement' and values of 0.5 to 0.7 are considered to indicate 'moderate agreement' (Schene et al, 2000).

**RESULTS**

Between 16th April and 30th June 2008 a total of 274 patients admitted to internal medicine wards were included in the study. The mean age was 65.3 years [standard deviation (SD) 17.69)]. A total of 137 (50.0%) were male and the majority (227; 82.8%) were admitted from home. Most were complex medical patients; the average number of co-morbidities was 3.3 (range 0 – 16); over one quarter (26.6%) were diabetics; and 105 (38.3%) were admitted with some type of infection, respiratory tract being the most common (40;14.6%). Thirty nine (14.2%) were not able to turn themselves independently. Calculating a Waterlow screening tool score for 74 patients was not possible because the body mass index (BMI) was unable to be assessed.

**Pressure ulcer incidence**

Fifteen participants (5.5) had an existing pressure ulcer and a further 12 (4.4%) developed a pressure ulcer during their hospital stay.

**Validity of the Waterlow screening tool**

Validity was assessed using the 200 patients for whom a Waterlow score was available. When the Waterlow screening tool score was calculated, 45 (22.5%) patients scored >15 indicating they were at risk of developing a pressure ulcer. Of these, six (13.3%) did so, compared to 3 of 155 (1.9%) who were not at risk ($p = 0.005$); sensitivity 0.67, (95% CI: 0.35 to 0.88); specificity 0.79, (95% CI: 0.73 to 0.85); PPV 0.13, (95% CI: 0.07 to 0.24); NPV 0.98, (95% CI: 0.94 to 0.99).

**Validity of 'ability to turn unaided'**

An additional validity analysis was conducted, on the total sample of 274 patients, using inability to turn unaided as the risk factor. Of the 39 patients who were unable to turn unaided, six developed a pressure ulcer (15.4%) compared with six of 235 (2.6%) who were able to turn unaided ($p = 0.003$). The positive predictive value of this risk factor (0.15) was higher than the Waterlow screening tool.

**Predictors for pressure ulcers**

Crude odds ratios and 95% confidence intervals for risk factors $p \leq 0.5$ are shown in Table 1. The very wide confidence intervals for many of the factors indicate a high degree of imprecision in these results; probably due to the low event rate.

**Interrater reliability.**

The interclass correlation for 'staging' pressure ulcers was .78 and inter-rater reliability for the Waterlow screening tool scoring was 1, indicating substantial agreement between raters.

## DISCUSSION

Limited evidence exists about the predictive efficacy of the Waterlow screening tool in acute hospital settings. Our results were consistent with those reported in a recent, comprehensive review of risk assessment scales (Pancorbo-Hidalgo et al, 2006). That is, in patients admitted to an acute hospital setting, the Waterlow screening tool is an inadequate strategy, by itself, for identifying those who may be at risk of developing a pressure ulcer. Although the Waterlow was effective in identifying those who would remain ulcer free, the tool as a whole was unable to discriminate well between those who later did or did not develop an ulcer.

The ethics of continuing to use a screen with limited predictive value requires discussion. It may be argued that it is preferable to screen, even though this may result in over-prediction, than not to screen at all, with the chance of missing some patients who may be at risk. However, in our experience, screening has not led to any reduction in the incidence of pressure injury but, rather, has led to reliance on screening and to a de-skilling of nurses. Moreover, routine screening for many other risks, such as falls, delirium and nutrition have become commonplace, consuming considerable nursing resources with little known benefit. If nursing is to be truly research based, we should be guided by evidence of true effectiveness and also take costs into account.

High false positive rates, such as those represented here have significant resource implications for organisations. For example, the policy at our own hospital recommends that everyone scoring >10 should be nursed on an overlay or specialist

foam mattress. The recommendation for those scoring >15 is for alternating pressure overlays or mattresses. Fluidised bead, low air loss and alternating pressure mattresses are required for those scoring 20+. In this institution, if these recommendations had been followed, almost 60% of all patients would have required some form of special mattress. This is clearly unsustainable and unnecessary and more accurate methods for identifying and managing those at high risk must be explored.

Six of the 45 patients, who were identified as being 'at risk' using the Waterlow scale, did develop a pressure injury. These patients were similar to others in this category for all of the variables we measured. Documentation in the medical record also indicated that they received at least 4$^{th}$ hourly pressure area care but we have no way of verifying that this occurred. All spent at least some of their hospital stay on a standard mattress but this was not different to other patients who scored in the 'at risk' range. The hospital has a limited number of static air overlay and alternate overlay mattresses, which are generally allocated to those with an already developed pressure injury.

Formatted: Superscript

A feature most likely to affect the accuracy of a screening tool is the incidence of the problem in the population. The incidence in this study was less than 10 percent so the positive predictive value was quite low. To show this relationship more clearly, the incidence of pressure ulcers and the respective positive predictive values from this study and from previously published work has been plotted (Figure 1). As can be seen in the graph, as the incidence of pressure ulcers rises, so does the positive predictive

value. Consequently, the Waterlow screening tool is probably of little value in acute care settings where the incidence of pressure ulcers is low relative to other settings.

We have speculated about other reasons for the poor predictive rates of the Waterlow screening tool and offer some insights. It became clear to the research nurses collecting data for the study that the BMI, an important component of the Waterlow screening tool, was rarely calculated. Nurses admitted that they frequently took a 'best guess', just so they could complete the scale. Some of the reasons for this include: the absence of scales for weighing in the ward area; patients not knowing their own weight or height; patients overstating their height (forgetting that they have become shorter with age); and an inability to get patients out of bed to estimate their height and weight.

Similarly, if nurses strongly believed that a patient needed a 'top of the range' mattress, they manipulated the Waterlow screening tool to meet the requirements needed for such a mattress to be issued. In this way, nurses were actually using their clinical judgement to identify patients at high risk, rather than depending on an arbitrary cut off score. Unfortunately, there tended to be a concomitant belief that as long as the patient was placed on the 'right' mattress, that other pressure area care, such as mobilisation was unnecessary.

Finally, scoring the Waterlow is complicated (Figure 2). Even among the trained research nurses there were frequent debates over the meaning of items and how a patient should be classified. For example, when asked to use the Waterlow screening tool to score the case study, there was100 percent agreement between the research

nurses, in terms of the broad classifications of the Waterlow screening tool, however, individual raw scores differed between all of the nurses. Moreover, in acute hospital settings, it is not unusual for daily changes in mobility and nutrition status to occur during a single episode of care. Consequently keeping pace with such changes is almost impossible; making the initial Waterlow score meaningless in many cases.

Other tools for assessing pressure injury risk such as the Norton score (Norton, 1996) and the Braden scale (Flanagan, 1993) are available but they also have limitations. Problems with these measures have been described elsewhere but they largely relate to the complex nature of the scoring and ambiguity associated with interpreting some items. (Papanikolaou, 2007). Sharp (2006) strongly argues that the only predeterminate for pressure injury is the inability of an individual to reposition without assistance (Sharpe and MacLaws, 2006). There is only one instrument that uses this single item for assessment; the Ramstadius tool (Ramstadius, 2000), which is not in wide use and still requires assessment as an intervention to prevent pressure ulcers. One limitation of all of the tools is that they detract from clinical judgement. An arbritary cut-off score cannot replace a thorough nursing assessment, careful and frequent ongoing observation and the development of individualised care plans.

**LIMITATIONS**

In some cases it was not possible to directly view all of the patient's pressure points. In these cases, information about the state of the patient's skin was obtained by asking the nurse caring for the patient or by extracting information from the medical record or nursing care plan. There is a possibility that this may have led to an underestimation of the true pressure ulcer incidence.

Although no changes were made to routine care during the study period, if research staff became aware of a new pressure ulcer, which had not been documented in the patients care plan or their medical record, staff caring for the patient were notified. Ethically we could not have done otherwise. However, we do not believe that our results were compromised. We were interested only in the ability of the Waterlow scale to predict which patients would develop a pressure ulcer; advising staff of the patients condition did not affect this.

**CONCLUSION**

Our study provides further evidence of the poor predictive validity of the Waterlow screening tool. Risk factors in our hospital were mostly related to mobility and nutrition. Consequently, efforts to ensure that patients are adequately nourished and are frequently mobilised while in hospital may offer greater benefit than screening, which may provide a false sense of security in the belief that 'something is being done'. We agree with findings from a recent Cochrane review (Moore and Cowman, 2008) that a suitably powered randomised controlled trial is urgently needed to provide definitive evidence about the usefulness of the Waterlow screening tool and to 'put it to bed' if it is shown to be no better than clinical judgement.

**ACKNOWLEDGEMENTS**

The study was internally funded by the hospital's Nursing Service. Matthew Walker, James Chambers, Becky Martin, Lynne McLaughlin and Debbie Anderson assisted with data collection.

None of the authors has any conflict of interest associated with this study.

**REFERENCES**


Black J, Baharestani MM, Cuddigan J, Dorner B, Edsberg L, Langemo D, Posthauer ME, Ratliff C, Taler, G (2007) National Pressure Ulcer Advisory Panel's updated pressure ulcer staging system. *Adv Skin Wound Care* **20**(5):269-274.

Bravo G, Potvin L (1991) Estimating the reliability of continuous measures with Cronbach's alpha or the intraclass correlation coefficient: toward the integration of two traditions. *J Clin Epidemiol* **44**(4-5):381-390.

Graves N, Birrell FA, Whitby M (2005) Modeling the economic losses from pressure ulcers among hospitalized patients in Australia. *Wound Repair Regen* **13**(5):462-467.

Keong N, Ricketts D, Alakeson N, Rust P (2004) Pressure sores following elective total hip arthroplasty: pitfalls of misinterpretation. *Ann R Coll Surg Engl* **86**(3): 174-176.

Kuwahara M, Tada H, Mashiba K, Yurugi S, Iioka H, Niitsuma K, Yasuda Y (2005) Mortality and recurrence rate after pressure ulcer operation for elderly long-term bedridden patients. *Ann Plast Surg,* **54**(6):629-632.

Landi F, Onder G, Russo A, Bernabei R (2007) Pressure ulcer and mortality in frail elderly people living in community. *Arch Gerontol Geriatr* **44**(Suppl 1):217-223.

Lapsley HM, Vogels R (1996) Cost and prevention of pressure ulcers in an acute teaching hospital. *Int J Qual Health Care,* **8**(1):61-66.

Moore ZE, COWMAN S (2008) Risk assessment tools for the prevention of pressure ulcers. *Cochrane Database Syst Rev***,** CD006471.

Nixon J, Cranny G, Bond S (2007) Skin alterations of intact skin and risk factors associated with pressure ulcer development in surgical patients: a cohort study. *Int J Nurs Stud* **44**(5):655-663.

Pancorbo-Hidalgo PL, Garcia-Fernandez FP, Lopez-Medina IM, Alvarez-Nieto C (2006) Risk assessment scales for pressure ulcer prevention: a systematic review. *J Adv Nurs* **54**(1):94-110.

Papanikolaou P, Lyne PA, Anthony D (2007) Risk assessment scales for pressure ulcers: a methodological review. *Int J Nurs Stud* **44**(2):285-296.

Ramstadius B (2000) Preventing institution acquired pressure ulcers. *ANJ* **7**:10

Schene AH, Koeter M, Van Wijngaarden B, Knudsen HC, Leese M, Ruggeri M, White IR, Vazquez-Barquero JL (2000) Methodology of a multi-site reliability study. EPSILON Study 3. European Psychiatric Services: Inputs Linked to Outcome Domains and Needs. *Br J Psychiatry Suppl* (39):s15-20.

Sharp CA, McLaws ML (2006) Estimating the risk of pressure ulcer development: is it truly evidence based? *Int Wound J* **3**(4):344-53.

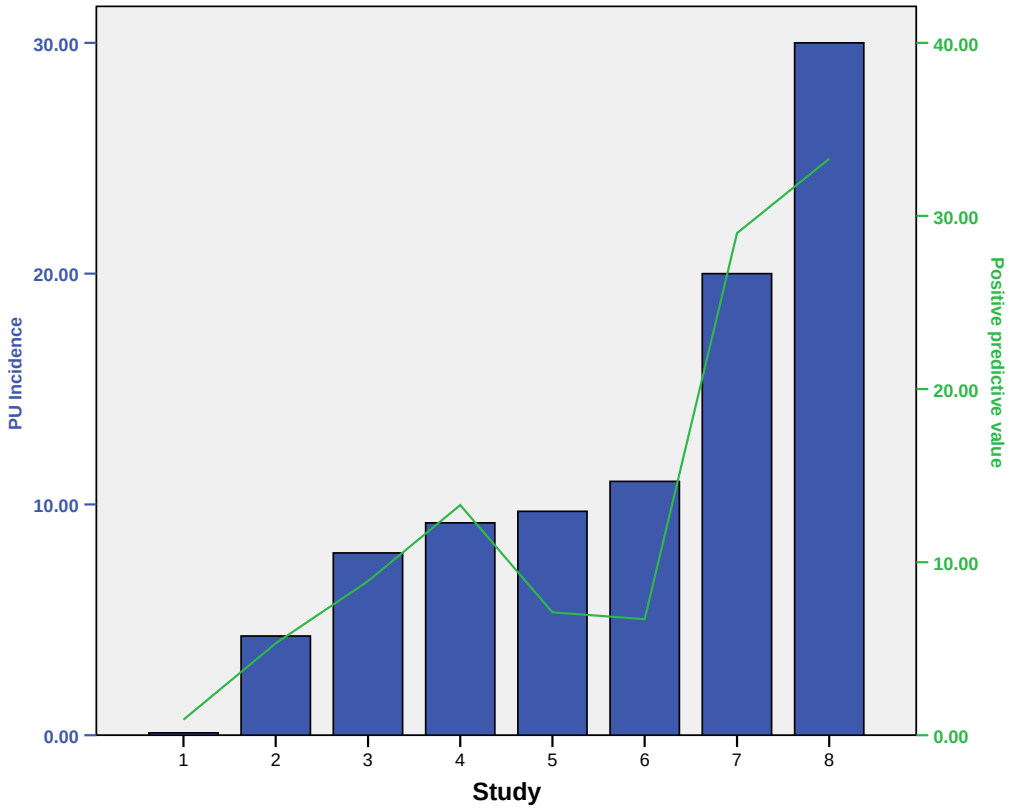
Waterlow, J. (1985) Pressure sores: a risk assessment card. *Nurs Times* **81**(48): 50-55.

Table 1. Factors associated with developing a pressure injury among a cohort of medical patients in an acute hospital setting

| Risk factors | Univariate analysis | | |
|---|---|---|---|
| | Crude OR* | 95% CI† | *P*‡ |
| Patient admitted from other hospital or nursing home | 0.34 | 0.14, 0.87 | 0.03 |
| Immobile before admission | 3.79 | 1.23, 11.46 | 0.03 |
| Able to turn unaided | 0.21 | 0.084, 0.53 | 0.002 |
| Waterlow score > 15 | 5.37 | 1.76, 16.42 | 0.004 |
| Received dietetic review | 10.35 | 4.15, 25.81 | 0.000 |
| Skin infection (including non-decubitus ulcer) | 4.08 | 1.45, 11.50 | 0.01 |
| Potassium level out of range | 3.08 | 1.091, 8.72 | 0.02 |
| Ability to transfer | 4.98 | 2.07, 11.98 | 0.000 |
| Ability to mobilise | 3.93 | 1.54, 10.05 | 0.007 |
| | MD | | |
| Older age | 10.17 | 2.68, 17.68 | 0.008 |
| Number of days of intravenous fluids | 0.51 | 0.28, 0.72 | < 0.000 |
| Number of days ‘nil by mouth’ | 0.77 | 0.32, 1.23 | 0.001 |

* Odds Ratio; † Confidence intervals; ‡ Level of significance

**Figure 1: Relationship between incidence and positive predictive values in eight reports of the validity of the Waterlow scale**




Data for studies 1,2,3,5,6,7,8 are taken from Pancorbo-Hidalgo et al [8]. Study 5 is data from the current study.

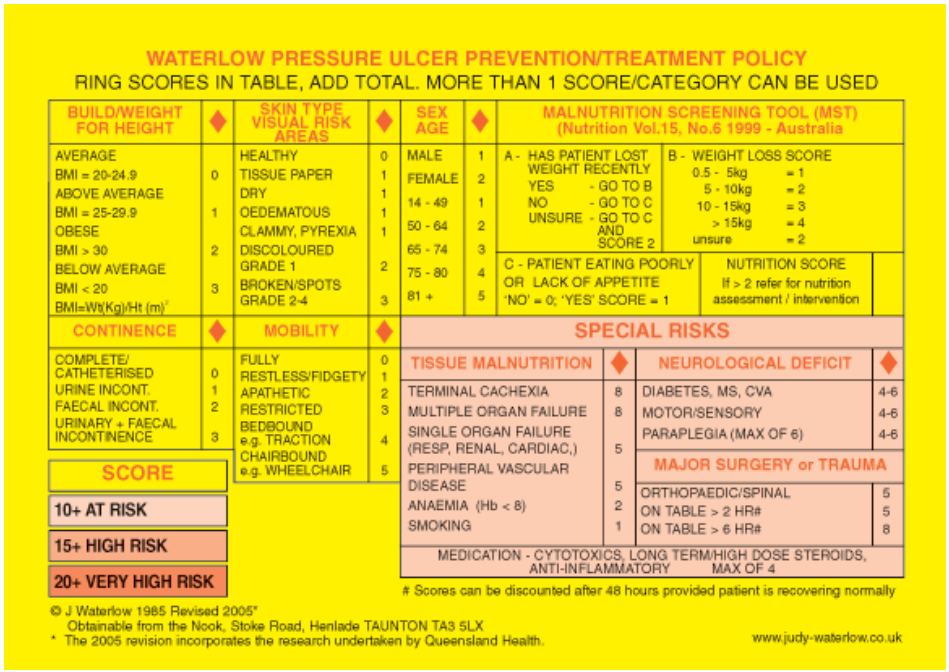



Figure 2: Waterlow pressure ulcer scale